

Bootstrap 기법을 이용한 서울지점 강우자료의 정량적 동질성 분석

A Quantative Homogeneity Analysis of Seoul Rainfall using Bootstrap

황석환*, 김중훈**, 유철상***, 정성원****, 유도근*****

Hwang, Seok Hwan, Kim, Joong Hoon, Yoo, Chulsang, Jung, Sung Won, Yoo Doguen

요 지

본 연구에서는 부트스트랩(Bootstrap) 기법을 이용하여 측우기 강우량 관측계열(CWK)과 근대우량계 강우량 관측계열(MRG)에 대해 동질성 분석을 실시하였다. 서로다른 두 자료계열에 대한 전통적인 통계적 동질성 검정 방법은 모집단의 분포형을 알고 있어야 검정결과가 유효하였기 때문에 모집단의 분포가 복잡한 기상자료들은 이러한 전통적 방법을 사용하여 동질성을 파악하는 것이 매우 어려웠고 결과로 제시된 통계적 유의성에 대해서도 의심의 여지가 있었다. 이러한 이유로 본 논문에서는 모집단을 가정하지 않아도 되는 비모수적 모의 방법인 부트스트랩 기법을 이용하여 두 자료계열간의 동질성 검정을 실시하였다. 분석 결과 M20의 CWK와 MRG는 미소한 기후의 경년변화(Trend)의 영향을 제외하면 동질성을 가진 자료로 볼 수 있었으나, 갈수기의 경우는 월강우량의 크기에 변화가 있으며 호우기의 경우는 일강우량의 크기 및 호우의 형태에 변화가 있는 것으로 나타났다.

핵심용어 : 부트스트랩, 기후변화, 측우기

1. 서론

장기간의 강수량기록은 기후 변동성 연구를 위한 귀중한 자료이다. 한반도에서는 정량적인 측우사업이 측우기의 발명과 더불어 15세기 중반부터 시작되었다. 이는 세계최장의 기록이나 우리나라 기후변화에 관련한 연구에서 강수량 관련 분석들은 1908년 이후 근대 관측자료가 대부분 이용되었다. 이는 측우기 자료가 현대의 일정한 계측시간 중심의 정량적 자료가 아닌 호우사상 중심의 강우지속기간과 지속기간 내 강수량형태로 기록되어 근대의 강수기록 기준과 다소 상이하여 근대 강수기록으로 변환시 다소의 불확실성이 증가한다는 이유에서 이다. 조선시대 강수량까지 포함한 서울의 강수량을 분석한 이병설(1970)은 1770년부터 1907년까지의 138년 동안의 연평균 강수량이 1908년에서 1960년까지의 연평균 강수량과 큰 차이를 보이지만 여름철만을 분석한 경우는 두 기간의 강수량 차이가 작음을 지적하였다. Arakawa(1956)는 1770-1907년의 측우기 강수량 자료와 서울 측후소에서 측정한 1908-1944년 동안의 강수량 자료를 비교 분석하여 측우기 자료가 상당한 신뢰성을 가진다고 하였다. Lim과 Jung(1992)은 서울 지역의 근대우량계 관측 연강수량은 측우기로 관측된 과거 조선시대의 연강수량과 동일한 수준임을 통계분석을 통해 제시한 바 있다. 그러나 기존의 통계적인 유의수준에서의 측우기 관측 강우량 자료계열과 근대강우량 관측 자료계열간의 동질성 분석은 해당 강우량 계열의 모집단 분포형에 대한 정확한 추정이 수반되어야만 가설검정에 대한 신뢰도가 높아진다는 단점이 있다. 따라서 본 연구에서는 모집단의 분포형을 가정하지 않는 재표본에 의한 모의 방법인 Bootstrap 기법을 이용하여 측우기 관측 강우량 자료계열과 근대우량계 강우량 관측계열의 평균과 분산에 대한 동질성 분석을 수행하였다. Pereira와 Oliveira(1984)는 연유출량 다지점 분해모형의 잔차항을 발생시키기 위해 부트스트랩 기법을 이용했다. Cover와 Unny(1986)는 유출량 ARMA 모형의 매개변수 불확실도를 추정하기 위해 부트스트랩 방법을 이용하였다. Sharma(1977)은 장기간의 유출량자료를 모의발생시키고자 부트스트랩 기법을 이용한 바 있다. Gary와 Paul(1997)은 가뭄빈도해석을 실시하고자 다지점 PAR(1) 모형에 부트스트랩 기법을 적용하

* 한국건설기술연구원 수자원연구실 연구원

Researcher, Korea Institute of Construction Technology, GyeongGi-Do 411-712, Korea
(e-mail: sukany@kict.re.kr)

** 고려대학교 공과대학 건축·사회환경공학부 교수

Prof., School of Architecture, Civil & Environmental Eng., Korea University, Seoul, 136-713, Korea
(e-mail: jaykim@korea.ac.kr)

*** 고려대학교 공과대학 건축·사회환경공학부 교수

Prof., School of Architecture, Civil & Environmental Eng., Korea University, Seoul, 136-713, Korea
(e-mail: envchul@korea.ac.kr)

**** 한국건설기술연구원 수자원연구실 책임연구원

Chief Researcher, Korea Institute of Construction Technology, GyeongGi-Do 411-712, Korea
(e-mail: swjung@kict.re.kr)

***** 고려대학교 공과대학 건축·사회환경공학부 박사과정

Doctoral Student, School of Architecture, Civil, and Environmental Eng., Korea University, Seoul, 136-713, Korea
(e-mail: godqhr425@korea.ac.kr)

였다. 김병식 등(2002)은 추계학적 모형의 잔차값을 부트스트랩 기법을 이용하여 연 및 월 하천 유출량 자료를 모의 복원 추출하였고 이를 Monte Carlo 모형, AR(1), PAR(1)과 같은 추계학적 방법과 비교하였을 때 기존의 방법과 부트스트랩 방법 모두 평균, 표준편차, 자기상관성은 잘 재현하였으나 왜곡도 계수의 경우는 부트스트랩이 기존의 방법보다 더 잘 재현한다고 하였다. 성기원(2003)은 모수적 부트스트랩 방법을 이용하여 단위도의 신뢰구간을 산정한 바 있다. 유철상 등(2007)은 부트스트랩 기법을 이용하여 유역 및 기상상태를 고려한 Clark 단위도의 매개변수 신뢰도 평가를 실시하여 집중시간 및 저류상수의 신뢰구간의 폭이 기존에 비해 좁다고 결론지었고 신뢰도 있는 집중시간 및 저류상수의 추정을 위해서는 최소 20개 이상의 독립된 호우사상이 필요하다고 하였다. 이러한 연구동향을 살펴보면 부트스트랩은 주로 실측자료에 근거한 장기간의 모의자료를 확보하기 위해 기존의 추계학적 방법의 대안으로 사용되었거나, 모형의 매개변수의 불확실도(혹은 신뢰구간)를 추정하는 연구에 주로 사용되어 왔음을 알 수 있다. 본 연구에서는 부트스트랩 방법을 이용하여 측우기 자료계열과 근대우량계 자료계열의 월별 동질성을 분석하였다. 이를 위해 부트스트랩 방법의 적용을 위해 필요한 통계적 독립성을 자기상관함수 분석, 연검정, 전환점 검정, 1차 자기상관계수 검정과 같은 기존의 통계적 방법을 이용하여 검정하였고, 통계적 분포의 동질성을 검정하기 위해 이표본 K-S 검정을 실시하였다. 더불어 월별 상자그림 분석을 통해 각 자료계열별로 통계치의 정량적 차이도 함께 비교하였다. 이를 통해 측우기 자료계열과 근대우량계 자료계열의 통계적 독립성과 동질성을 확인한 후 부트스트랩 방법을 이용하여 관측자료로부터 ECDF(경험누가확률분포)를 산정하여 실측자료에 기반한 두 자료계열의 정량적 동질성을 분석하여 보았다.

2. 분석자료

2.1 월별 동질성 분석을 위한 통계치의 설정

다음은 측우기 관측계열과 근대우량계 관측계열간의 월별 동질성 분석을 하기 위해 선정한 4개의 통계치로 강우변화 특성을 효과적으로 나타내 줄 수 있는 특성들이다. 강우는 강우량과 같은 양적 특성과 발생횟수와 같은 빈도 특성 그리고 지속시간에 따른 강도 특성을 복합적으로 가지고 있기 때문에 강우의 변화를 파악하기 위해서는 양적 통계치는 물론 빈도와 강도를 적절히 표현할 수 있는 통계치를 동시에 비교해야만 강우특성의 변화여부에 대한 적절한 판단이 가능하다. 표1의 *Monthly rainfall*은 4월에서 10월까지 각 월의 월강우량이고 $D_{max} ratio$ 는 월강우량에 대한 해당월 일최대 강우량이 차지하는 비율의 변화를 분석하기 위한 통계치이다. $N_{rainy\ days}$ 는 각 월별 평균적인 강우일수를 산정한 통계치이고 $I_{rainy\ days}$ 는 각 월별 강우일수에 대한 월강우량의 비로 강우일에 대한 평균강우강도이다. *Monthly rainfall*은 강우량의 정량적인 변화를 파악하기 위한 기본 통계치이고 $D_{max} ratio$ 는 일최대 강우량의 정량적 변화를 분석하여 단기 강우사상의 규모 및 특성 변화를 파악할 수 있기 때문에 선택하였다. $N_{rainy\ days}$ 는 강우일수의 변화를 정량적으로 나타내므로 이를 분석하면 강우발생 경향의 변화 여부를 쉽게 파악할 수 있고 $I_{rainy\ days}$ 는 강우강도의 변화를 정량적으로 분석 가능하기 때문에 사용하였다.

3. 부트스트랩 기본이론

3.1 부트스트랩(Bootstrap) 기본 원리

일반적으로 강우를 포함한 기상인자의 정량화를 위하여 모수적인 방법을 사용하고 있으나 이를 극복할 필요가 있다. 이 모수적 방법은 기존의 통계적인 방법으로 익히 알려진 확률밀도함수에 특정 기상인자의 특성이 완벽하게 부합한다는 가정에 근거한다. 다시 말해 표준오차의 신뢰한계를 명확히 산정할 수 있어야 한다. 그러나 기상요소와 같이 변동성이 크거나 불규칙한 장기적인 거동특성의 경우에 그러한 가정이 적합하지 않을 가능성이 매우 크다. 그리고 기존의 방법으로 통계적 유의성을 검정할 때 모집단의 분포를 모르기 때문에 모집단을 정규분포로 가정하고 표본의 크기가 작은 경우 t -분포를 사용한다. 이는 모분산 대신 표본분산을 사용하는 방법으로 정규분포에 비해 t -분포는 더 납작한데 이는 모분산대신 표본분산을 사용하여 더 큰 불확실성을 갖게 되어 꼬리가 긴 분포를 만들기 때문이다. 그러나 여기서 t -분포 역시 표본의 크기에 대한 정규분포인 모집단과의 관계를 나타낼 뿐 실제로 모집단이 정규분포인지에 대한 의문은 여전히 남아있다. 따라서 관측자료가 가지고 있는 특성을 그대로 보전하여 재표본 방법에 의해 직접 분포형을 추정하는 자료모의 방법이 제시되었고, 이러한 재표본 방법 중 가장 간단하고 보편적인 방법으로 Efron(1979)의 부트스트랩(bootstrap)기법이 있다. Efron(1979)에 의해 제안된 부트스트랩기법은 주어진 표본에 근거하여 재표본(resampling)을 취하여 연구대상이 되는 통계량의 성질을 파악한다. 부트스트랩기법은 기존의 방법들과 달리 주어진 검정통계량의 분포를 가정하지 않고, 주어진 검정통계량의 분포를 직접 근사한다. 또한 부트스트랩기법은 근본적으로 모형을 설정하지 않는 자료에 기반한 모의방법이기 때문에 모형에 의존하지 않는다 라는 이점을 가지고 있다(전명식, 1996). 따라서 표준오차의 신뢰한계를 구할 수 없는 경우, 그 표본을 마치 모집단인 것처럼 다루는 재표본추

출 방법인 부트스트랩 분석의 기본 원리 및 절차는 매우 간단하다. 원래의 표본에서 재표본을 추출한 후 재표본에 대한 추정치를 산정하고 이 두 단계를 여러 번 반복하여 재표본 추정치의 산포를 파악하는 것이다. 표준 비모수적 부트스트랩의 기본 절차는, 주어진 확률표본 x_1, x_2, \dots, x_n 으로부터 복원 추출 방법으로 부트스트랩 표본 $x_1^*, x_2^*, \dots, x_n^*$ 을 얻는다. 이때 무작위 추출이므로 추출 가능한 부트스트랩 표본의 총 수는 n^n 임을 알 수 있다. 다음으로 이 n^n 개(대부분 이보다 작음)의 부트스트랩 표본에 대해 구하고자 하는 통계량 $\hat{\theta}^*$ 를 얻을 수 있으며 이를 통해 추정 통계량 $\hat{\theta}^*$ 의 부트스트랩 분포를 추정하게 된다. Efron과 Tibshirani(1986)은 정확한 신뢰구간 추정값을 계산하기 위해서는 최소한 1000개의 부트스트랩 표본이 적절하다고 제안하였다. 따라서 B 개(본 연구에서는 1000개)의 부트스트랩 표본으로부터 통계량 $\hat{\theta}^*$ 을 구하고 오름차순으로 정렬하여 θ 의 추정량($\hat{\theta}$)과 순번이 정해진 부트스트랩 추정값($\hat{\theta}_i^*$)을 산정한다. 이러한 부트스트랩 추정값을 이용하여 θ 에 대한 $100(1-\alpha)\%$ 부트스트랩 신뢰구간을 구할 수 있다. B 개의 부트스트랩 추정값 $\hat{\theta}_i^*$ 로부터 표본 평균과 표본 분산은 다음과 같다.

$$\bar{\theta}^* = \frac{1}{B} \sum_{i=1}^B \hat{\theta}_i^* \quad (1)$$

$$v_{\theta}^* = \frac{1}{B-1} \sum_{i=1}^B (\hat{\theta}_i^* - \bar{\theta}^*)^2 \quad (2)$$

여기서 v_{θ}^* 은 $\hat{\theta}$ 표본편차의 추정량이고 그 분포가 근사적으로 정규분포이면 θ 에 대한 $100(1-\alpha)\%$ 신뢰구간은 다음과 같다.

$$\hat{\theta} \pm z_{\alpha/2} v_{\theta}^* \quad (3)$$

여기서 z_{α} 는 표준정규분포의 상위 $100 \times \alpha\%$ 백분위수이다.

3.2 평균과 분산변화에 대한 부트스트랩 유의성 검정

앞서 언급한 바와 같이 측우기의 관측최소 단위가 1분(1分=2mm)이므로 0.1mm단위로 측정하는 근대 우량계 관측과 비교하면 소량의 강수는 측정되지 않았을 가능성이 높다(정현숙, 1994) 라고 보고 있다. 상식적으로 보아도 당시의 관측수준이 현대보다 높다고 보긴 힘들다. 그러나 이러한 연구결과는 미소한 수준의 관측 정확도가 높고 낮음에 대한 언급이지, 본 논문의 분석 초점인 통계학적으로 유의성을 가질 만큼 정량적인 차이가 있고 없음에 대한 판단 결과는 아니다. 즉, 호우사상의 크기나 빈도 등의 변화량에 따라 판단하는 경년변동, 기후변화 등을 파악하는 목적에서는 미소한 수준의 정량적인 차이가 중요 하긴 하지만, 여름철을 기준으로, 분석하고자 하는 목적에 비추어 비교대상간의 차이에 유의성이 없다면 두 자료를 활용하는 데는 큰 문제가 되지 않을 것이다. 본 논문에서는 부트스트랩 방법을 이용하여 실측 자료에 근거한 측우기 자료와 근대 우량계 자료의 *Monthly rainfall*, *D_{max} ratio*, *N_{rainy days}*, *I_{rainy days}*에 대한 평균과 분산의 차이에 대한 유의성 검정을 실시하였다. 측우기 자료의 경우 강설이 고려되지 않았기 때문에 4월에서 10월까지의 연단위 강수량 자료와 각 월별 자료에 대한 검정을 실시하였다. 더불어 이러한 미소수준의 차이를 줄여 측우기 관측단위와 동일한 조건에서 유의성 검정을 실시하기 위하여 근대 우량계 자료중 2mm이하의 경우는 무강우로 간주하고 동일한 유의성 검정을 실시하였다. 이는 두가지 목적에서 실시되었는데 첫째는 측우기 자료의 정확도를 동일한 수준에서 가늠해 보는 것이며, 두 번째는 연단위 자료의 특성이 유사한 경우 월별 강우의 경년변동(trend) 특성을 파악해 보고자 하는 목적이다. 평균에 대한 귀무가설(null hypothesis)은 측우기 자료와 근대 우량계 자료의 "평균은 같다"이고, 대립가설의 경우는 측우기 자료와 근대 우량계 자료는 "평균이 같지 않다"라고 정의하였다. 즉, 양측검정에 해당한다. 분산에 대한 귀무가설도 측우기 자료와 근대 우량계 자료의 "분산은 같다"이고, 대립가설의 경우는 측우기 자료가 근대 우량계 자료와 "분산이 같지 않다"라고 정의하였고, 이경우도 양측검정에 해당한다. 그리고 부트스트랩을 이용하여 자료를 재생성하기 위해 비모수적인 표준 부트스트랩 방법을 이용하였고 각 자료계열별로 표본추출회수 B 는 1000회를 실시하였다.

4. 분석결과

본 논문에서 측우기와 근대우량계 자료계열간의 변동특성을 파악하기 위해 제시한 4가지 통계특성인 *Monthly rainfall*, *D_{max} ratio*, *N_{rainy days}*, *I_{rainy days}*에 대한 월별 평균과 분산의 차이에 대한 표준 부트스트랩 유의성 검정 결과는 다음 표1과 같다. 자료의 분석대상 기간은 4월에서 10월까지로 월별로 95%신뢰수준(5% 신뢰한계) 검정을 실시하였다. *Monthly rainfall*의 경우 M00은 4월을 제외하면 평균차에 대한 유의성은 보이지 않았고 분산은 분석대상 기간 전체에서 차이에 대한 유의성을 보이지 않았다. M20은 분석대상기간 전체에서평균과 분산의 차이에 대한 유의성을 보이지 않았다. 이를 통해 볼 때 두계열간의

강우의 양적 차이는 크지 않으며 특히 2mm 이하의 강우가 차지하는 비중이 크지 않음을 알 수 있다.

Table 1. Results of hypothesis test using bootstrap methods (two-sides test)

Statistics	Groups	Month	Mean					Variance					
			p	C.L.	t_{c_L}	t_{c_U}	H_0	p	C.L.	t_{c_L}	t_{c_U}	H_0	
<i>Monthly rainfall</i>	M00	APR	0.022	0.05	-28.23	-1.40	R	0.112	0.05	-3269	194	A	
		MAY	0.146	0.05	-24.13	3.95	A	0.768	0.05	-1613	1979	A	
		JUN	0.312	0.05	-42.08	11.56	A	0.300	0.05	-10072	2460	A	
		JUL	0.176	0.05	-93.03	19.34	A	0.804	0.05	-29605	34035	A	
		AUG	0.344	0.05	-70.94	23.41	A	0.396	0.05	-36766	10103	A	
		SEP	0.062	0.05	-53.43	1.95	A	0.996	0.05	-7266	9952	A	
	M20	OCT	0.442	0.05	-14.48	5.87	A	0.456	0.05	-669	1222	A	
		APR	0.080	0.05	-26.01	1.40	A	0.112	0.05	-3150	216	A	
		MAY	0.284	0.05	-22.27	7.52	A	0.734	0.05	-1540	2105	A	
		JUN	0.436	0.05	-37.53	16.30	A	0.364	0.05	-9613	2690	A	
		JUL	0.206	0.05	-92.42	19.33	A	0.766	0.05	-28047	34849	A	
		AUG	0.394	0.05	-76.75	25.03	A	0.346	0.05	-35830	9915	A	
	<i>D_{max} ratio</i>	M00	SEP	0.086	0.05	-51.07	4.31	A	0.998	0.05	-8182	10299	A
			OCT	0.730	0.05	-12.07	8.41	A	0.516	0.05	-620	1245	A
			APR	0.024	0.05	0.007	0.108	R	0.122	0.05	-0.003	0.025	A
			MAY	0.038	0.05	0.002	0.082	R	0.018	0.05	0.002	0.023	R
JUN			0.888	0.05	-0.039	0.043	A	0.506	0.05	-0.007	0.015	A	
JUL			0.964	0.05	-0.027	0.029	A	0.002	0.05	0.003	0.014	R	
M20		AUG	0.762	0.05	-0.031	0.036	A	0.080	0.05	-0.001	0.015	A	
		SEP	0.830	0.05	-0.055	0.043	A	0.020	0.05	0.003	0.033	R	
		OCT	0.116	0.05	-0.007	0.100	A	0.002	0.05	0.011	0.045	R	
		APR	0.326	0.05	-0.026	0.081	A	0.536	0.05	-0.011	0.020	A	
		MAY	0.160	0.05	-0.012	0.071	A	0.162	0.05	-0.004	0.020	A	
		JUN	0.590	0.05	-0.053	0.029	A	0.838	0.05	-0.011	0.013	A	
<i>N_{rainy days}</i>		M00	JUL	0.850	0.05	-0.032	0.025	A	0.002	0.05	0.003	0.013	R
			AUG	0.964	0.05	-0.038	0.031	A	0.246	0.05	-0.003	0.014	A
			SEP	0.264	0.05	-0.082	0.022	A	0.332	0.05	-0.009	0.026	A
			OCT	0.702	0.05	-0.052	0.074	A	0.262	0.05	-0.009	0.031	A
	APR		0.000	0.05	-3.76	-2.32	R	0.112	0.05	-5.56	0.38	A	
	MAY		0.000	0.05	-3.93	-2.47	R	0.126	0.05	-5.32	0.57	A	
	M20	JUN	0.000	0.05	-4.40	-2.77	R	0.712	0.05	-3.74	2.62	A	
		JUL	0.000	0.05	-5.45	-3.22	R	0.008	0.05	2.34	15.20	R	
		AUG	0.000	0.05	-4.09	-1.99	R	0.702	0.05	-4.50	6.90	A	
		SEP	0.000	0.05	-3.96	-2.38	R	0.940	0.05	-3.83	3.63	A	
		OCT	0.000	0.05	-3.33	-1.96	R	0.124	0.05	-4.65	0.46	A	
		APR	0.932	0.05	-0.73	0.61	A	0.902	0.05	-2.39	2.27	A	
	<i>I_{rainy days}</i>	M00	MAY	0.340	0.05	-0.32	0.93	A	0.102	0.05	-0.29	3.51	A
			JUN	0.688	0.05	-0.61	0.91	A	0.106	0.05	-0.52	5.13	A
			JUL	0.632	0.05	-0.82	1.28	A	0.002	0.05	3.96	16.96	R
			AUG	0.080	0.05	-0.12	1.80	A	0.040	0.05	0.22	10.12	R
SEP			0.984	0.05	-0.69	0.76	A	0.066	0.05	-0.17	6.28	A	
OCT			0.820	0.05	-0.50	0.62	A	0.084	0.05	-0.37	3.62	A	
M20		APR	0.000	0.05	2.09	4.93	R	0.002	0.05	9.06	40.89	R	
		MAY	0.000	0.05	2.27	5.57	R	0.008	0.05	6.63	53.01	R	
		JUN	0.002	0.05	2.51	7.27	R	0.008	0.05	13.81	97.10	R	
		JUL	0.002	0.05	1.62	6.87	R	0.104	0.05	-9.45	99.09	A	
		AUG	0.014	0.05	0.66	6.31	R	0.782	0.05	-68.25	44.37	A	
		SEP	0.036	0.05	0.13	5.92	R	0.184	0.05	-33.34	142.70	A	
M20		OCT	0.004	0.05	1.23	4.84	R	0.012	0.05	6.85	71.05	R	
		APR	0.118	0.05	-3.66	0.25	A	0.602	0.05	-52.39	20.50	A	
		MAY	0.028	0.05	-4.91	-0.28	R	0.120	0.05	-81.77	6.82	A	
		JUN	0.170	0.05	-4.34	0.97	A	0.284	0.05	-18.50	69.77	A	
	JUL	0.014	0.05	-6.99	-1.09	R	0.882	0.05	-48.74	61.86	A		
	AUG	0.004	0.05	-7.94	-1.33	R	0.012	0.05	-191.27	-13.79	R		
M20	SEP	0.012	0.05	-8.14	-0.72	R	0.422	0.05	-257.83	72.99	A		
	OCT	0.384	0.05	-2.85	1.04	A	0.330	0.05	-23.48	57.47	A		

D_{max} ratio의 경우 M00은 4월과 5월이 평균의 차이에 유의성을 보였고 5월, 7월, 9월, 10월에서 분산의 차이에 유의성을 보였다. 평균의 차이는 강우량이 적은 봄철에서 두드러 졌다. M20은 평균의 차이에는 전기간에서 유의성을 보이지 않았고 분산의 차이는 7월에서만 유의성을 보였다. 이를 통해 강우량이 가장 많은 7월의 단기호우 변동폭은 과거와 근대이후에 변화가 있었음을 추정할 수 있다. $N_{rainy days}$ 의 경우 M00은 전체기간에서 평균의 차이에 유의성을 보였고 분산은 7월에서만 유의성을 보였다. 이를 통해 볼 때 두자료계열간의 강우일수는 확연한 차이가 있음을 알 수 있고 이는 강우량 변화가 크지 않음을 고려해 볼 때 관측 정밀도가 가장 큰 요인일 것으로 판단된다. M20은 평균의 차이에는 전기간에서 유의성을 보이지 않았고 분산의 차이는 7월과 8월에서만 유의성을 보였다. 특히 7월의 p 값이 매우 작아 연별 강우일수의 변동폭에 변화가 있음을 알 수 있다. $I_{rainy days}$ 의 경우 M00은 전기간에서 평균의 차이에 유의성을 보였고 분산은 4월, 5월, 6월과 10월에서 유의성을 보였다. 특히 분산은 집중호우가 자주 발생하

는 7월, 8월, 9월을 제외한 기간에서 차이에 대한 유의성을 보이고 있는데 이는 강우량에 대한 민감도가 우기에 비해 상대적으로 갈수기인 4월, 5월, 6월, 10월이 크기 때문으로 판단된다. M20은 5월, 7월, 8월, 9월의 평균 차이에서 유의성을 보였고 분산은 8월의 경우만 그 차이에 유의성을 보였다. 이를 통해 볼 때 여름철 강우강도는 과거에 비해 근대이후 변화가 있음을 추정 수 있다.

5. 결론

본 장에서는 부트스트랩 기법을 이용하여 측우기 강우량 관측계열(CWK)과 근대우량계 강우량 관측계열(MRG) 사이의 동질성을 분석해 보았다. 부트스트랩을 이용하여 두 집단의 동질성 분석을 한 결과를 정리하면, 부트스트랩 기법을 이용하여 95% 신뢰구간에서 CWK와 MRG간의 동질성 검정결과(양측검정), M20의 경우 *Monthly rainfall*은 4월에서 9월 모두 평균과 분산의 차이에 유의성을 보이지 않았다. D_{max} ratio는, 모든 달에서 평균의 차이에 유의성을 보이지 않았고, 분산의 차이는 7월만이 유의성을 보였다. $N_{rainy\ days}$ 는 모든 달에서 평균의 차이에 유의성을 보이지 않았고, 분산의 차이는 7월과 8월만이 유의성을 보였다. $I_{rainy\ days}$ 는 5월, 7월, 8월, 9월이 평균의 차이에 유의성을 보였고, 분산은 8월의 경우만 차이에 유의성을 보였다. 본 연구를 통하여 *Monthly rainfall*의 경우는 상대적으로 갈수기의 변화가 두드러지고 있으며, D_{max} ratio나 $N_{rainy\ days}$ 의 경우는 7월(호우기)의 강우량의 분산에 변화를 보이고 있다. 이는 갈수기의 경우는 월강우량의 크기에 변화가 있으며, 호우기의 경우는 일강우량의 크기 및 호우의 형태에 변화가 있음을 나타낸다고 볼 수 있다. 이는 측우기 자료자체의 신뢰도에 문제일 수 있으나 기존 연구 및 본 연구에서 나타난 월별 일관성을 볼 때 단편적이지만 기후변화 양상을 나타내는 것으로 판단된다.

6. 참고문헌

- 김병식, 김형수, 서병하 (2002). "Bootstrap 방법에 의한 하천유출량 모의와 왜곡도." **한국수자원학회논문집**, 제35권 제3호, pp. 275-284.
- 성기원 (2003). "유역의 수문학적 상사성을 이용한 Nash 모형의 불확실성 평가." **한국수자원학회논문집**, 제36권, 제3호, pp. 399-411.
- 유철상, 이지호, 김기욱 (2007). "유역 및 기상상태를 고려한 Clark 단위도의 매개변수 평가: 2. 매개변수의 변동성 추정." **한국수자원학회논문집**, 제40권, 제2호, pp. 171-182.
- 이병설 (1970). "서울의 연강수량 및 하기강수량의 Normality에 관한 연구." **한국기상학회지**, 5, 11-14
- Arakawa, H. (1956). "On the secular variation of annual total of rainfall at Seoul from 1770 to 1944." *Journal of the Korean Meteorological Society*, Vol. 7, No. 2, pp. 205-211.
- Cover, K.A., Unny, T.E. (1986). "Application of computer intensive statistics to parameter uncertainty in streamflow synthesis." *Water Resources. Research. Bull.*, Vol. 22, No. 3, pp 495-507.
- Efron, B. (1979). "Bootstrap method: another look at the jackknife." *Journal of Statistics*, Vol. 7, pp. 1-26.
- Gary, D.T., and Dunne, P. (1997). "Bootstrap Position Analysis for forecasting Low Flow Frequency." *Journal of Water Resources Planning and Management*, Nov-Dev, pp. 355-367.
- Lim, G.H., and Jung, H.S. (1992). "Interannual variation of the annual precipitations at Seoul, 1771-1990." *Journal of the Korean Meteorological Society*, Vol. 28, pp. 125-132.
- Pereira, M.V.F., Oliveria, G.C., Costa, C.G., Kelman, J. (1984). "Stochastic stream flow models for hydroelectric system." *Water Resources. Research*. Vol. 20, No. 3, pp. 379-390.
- Sharma, A., Tarboton, D.G., and Lall, U. (1997). "Streamflow simulation: a non-parametric approach." *Water Resources. Research.*, Vol. 33, No. 2, pp.291-308.